TÓM LƯỢC BÀI GIẢNG

(Vũ Quốc Hoàng)

KIỂU CHUỖI string (C++)

Chủ đề

- Kiểu chuỗi `string` (C++)

Tài liệu

[1] Tony Gaddis, *Starting out with C++ From Control Structures through Objects*, Pearson, 8th edition, 2015.

[2] Vũ Quốc Hoàng, *Bí kíp luyện Lập trình C (Quyển 1)*, hBook, 2017.

Đọc tài liệu

- Kiểu chuỗi `string` (C++): Section 2.8, 3.8, 10.7 [1]

Kiến thức

- Đoạn mã C++ minh họa 1

```
#include <iostream>
#include <string>
using namespace std;

int main()
{
    string firstname, fullname;

    cout << "What's your first name? ";
    cin >> firstname;
    cout << "Hi " << firstname << endl;

    cout << "What's your full name? ";
    fflush(stdin);
    getline(cin, fullname);
    cout << "Hello " << fullname << endl;
    cout << "Hello " + fullname << endl;
    printf("Hello %s\n", fullname.c_str());

    return 0;
}
```

- Thư viện chuẩn C++ cung cấp một kiểu dữ liệu để lưu trữ và thao tác với chuỗi là kiểu `string`. Kiểu này cùng với các hàm thao tác trên nó được để trong không gian tên `std` và

được #include từ file tiêu đề string. Về mặt thuật ngữ, kiểu string là một *kiểu lớp* (class). Tuy nhiên ta sẽ được học về kiểu lớp (và các khái niệm khác của "*lập trình hướng đối tượng*") trong môn sau. Tạm thời ta có thể hiểu đó là một kiểu cấu trúc "đặc biệt".

- Các biến kiểu string được khai báo như các biến đã học và có thể được gán giá trị từ *hằng chuỗi* C (tức là dãy kí tự trong cặp nháy kép "…").
- Các biến kiểu string cũng có thể được nhập/xuất với cin/cout như các biến khác.
- Hàm getline (cũng trong không gian tên std) có thể được dùng để nhập chuỗi có khoảng trắng cho biến kiểu string.
- Cũng như trong C, khi nhập chuỗi (và kí tự) ta nên xóa *bộ đệm nhập* (input buffer) trước. Điều này có thể được thực hiện bằng lời gọi hàm: fflush(stdin).
- Kiểu string cung cấp vài thao tác trên chuỗi được thực hiện bằng các toán tử, mà quan trọng nhất là toán tử *nối chuỗi* (concatenate) +. Toán tử này giúp nối chuỗi trong hai biến string hay chuỗi trong biến string với chuỗi kiểu C (mảng kí tự kết thúc bởi kí tự NULL mà ta đã học). Lưu ý là toán tử + kí hiệu cho nhiều thao tác khác nhau (cộng 2 số, nối 2 chuỗi, …). Điều này được gọi là *nạp chồng toán tử* (operator overloading) trong C++ mà ta sẽ học sau.
- Kiểu string cũng cung cấp nhiều *hàm thành phần* (member function), còn được gọi là *phương thức* (method), thao tác trên dữ liệu của kiểu (còn gọi là *đối tượng*, object), giúp cung cấp các chức năng xử lý chuỗi. *Toán tử truy cập thành phần* (toán tử .) giúp truy cập phương thức như cách nó được dùng để truy cập thành phần (dữ liệu) của kiểu cấu trúc.
- Phương thức c_str của kiểu string giúp "lấy về" chuỗi kiểu C tương ứng với chuỗi của đối tượng string. Điều này cũng cho thấy rằng string của C++ là kiểu "*bọc*" (wrapper) kiểu chuỗi của C.
- Đoạn mã C++ minh họa 2

```
#include <iostream>
#include <string>
using namespace std;

void split(string namelist, string names[], int &n)
{
    namelist += ',';
    n = 0;
    int pos = 0, i;
    while((i = namelist.find(',', pos)) >= 0)
    {
        string name = namelist.substr(pos, i - pos);
        if(name.length() > 0) // if(name != "")
        {
            name[0] = toupper(name[0]);
            names[n++] = name;
        }
        pos = i + 1;
    }
```

```
}

void sort(string names[], int n)
{
    for(int i = 0; i < n; i++)
        for(int j = i + 1; j < n; j++)
            if(names[i] > names[j])
            {
                string temp = names[i];
                names[i] = names[j];
                names[j] = temp;
            }
}

void output(string names[], int n)
{
    for(int i = 0; i < n; i++)
        cout << "Hello " << names[i] << endl;
}

int main()
{
    string namelist, names[100];
    int n;

    cout << "Enter a list of names separated by comma: ";
    getline(cin, namelist);

    split(namelist, names, n);
    sort(names, n);
    output(names, n);

    return 0;
}
```

- Một vài toán tử hay dùng khác trên kiểu `string` là: += (dùng để nối thêm vào chuỗi), [...] (dùng để truy cập kí tự của chuỗi), != , ==, >, ... (dùng để so sánh chuỗi theo thứ tự anphabet)
- Một vài phương thức hay dùng khác của kiểu `string` là: `length` (dùng để tính chiều dài chuỗi), `find` (dùng để tìm vị trí xuất hiện của một kí tự hay chuỗi con nào đó trên chuỗi), `substr` (dùng để trích chuỗi con từ chuỗi), ...
- Kiểu `string` cũng có thể được dùng với mảng hay truyền nhận đối số cho hàm hay là thành phần của kiểu cấu trúc như các kiểu khác.

Kĩ năng

- Thành thạo việc dùng kiểu `string` để xử lý chuỗi

- Thành thạo việc “tích hợp” và dùng chung kiểu `string` với các kiểu dữ liệu khác, nhất là kiểu chuỗi của C

Lưu ý

- Chuỗi là dạng dữ liệu rất quan trọng, hơn nữa kiểu `string` của C++ giúp việc xử lý chuỗi đơn giản và tiện lợi hơn nên sinh viên cần thành thạo việc sử dụng kiểu này

Bài tập

1. Tìm một nguồn (web) tin cậy và tra cứu kiểu `string` (mô tả kiểu, các toán tử, các phương thức, …) của C++.
2. Làm lại các bài tập trong Bài 4.4 [2] bằng kiểu `string` của C++.